THE USE OF SPECIFIED TCF TARGET GENES TO IDENTIFY DRUGS FOR THE TREATMENT OF CANCER, IN PARTICULAR COLORECTAL CANCER, IN WHICH TCF/β-CATENIN/WNT SIGNALLING PLAYS A CENTRAL ROLE

5   The present invention relates to the use of genes whose expression is regulated by TCF/β-catenin complexes in colon carcinoma cells, for the identification and development of small molecule inhibitors, antibodies, antisense molecules, RNA interference (RNAi) molecules and

10   gene therapies against these target genes and/or their expression product, for the treatment of cancer in which deregulated TCF/β-catenin signalling occurs, in particular colorectal cancer and melanomas. In addition the invention relates to a method for the development of the small

15   molecule inhibitors and antibodies. The invention also relates to the small molecule inhibitors, antibodies, antisense molecules, RNAi molecules and therapeutic genes per se and to their use in the treatment and diagnosis of cancer in which deregulated TCF/β-catenin/WNT signalling

20   occurs and to pharmaceutical compositions comprising them.

The colorectal mucosa contains large numbers of invaginations known as the crypts of Lieberkühn. Epithelial cells in these structures are constantly renewed in a coordinated series of events comprising proliferation, cell

25   migration and differentiation along the crypt axis towards the intestinal lumen. Pluripotent stem cells are believed to reside at the bottom positions of the crypt. From these stem cells, progenitors are generated that occupy the lower third of the crypt, the amplification compartment. Cells in

30   this compartment divide approximately every 12 hours until their migration brings them to the mid-crypt region. Here, they cease proliferating and differentiate into one of the functional cell types of the colon. At the surface

epithelium, cells undergo apoptosis and/or extrusion into the lumen. The complete process takes approximately 3-5 days.

Colorectal cancer (CRC) is one of the most common malignancies in the western world. The transition of an intestinal epithelial cell into a fully transformed, metastatic cancer cell is a slow process, requiring the accumulation of mutations in multiple proto-oncogenes and tumour suppressor genes. The *APC* gene, originally cloned from patients with the rare genetic disorder Familial Adenomatous Polyposis, is mutated in the vast majority of sporadic CRCs.

The APC protein resides in the so-called destruction complex, together with GSK3β, axin/conductin and β-catenin. In this complex, phosphorylation by GSK3β targets β-catenin for ubiquitination and destruction by the proteasome. Signalling by the extracellular factor WNT inhibits GSK3β activity. As a result, β-catenin accumulates in the nucleus where it binds members of the TCF family and converts these WNT effectors (from transcriptional repressors into transcriptional activators. The terms "TCF/β-catenin signalling" and "WNT-signalling" are commonly used to describe the same signalling pathway.

In cancer, truncating mutations in *APC* and axin/conductin, as well as mutations in the GSK3β-target residues in β-catenin all lead to the formation of constitutive nuclear β-catenin/TCF complexes. Activating mutations of the WNT pathway are the only known genetic alterations present in early premalignant lesions in the intestine, such as aberrant crypt foci and small polyps. Thus, these mutations appear to initiate the transformation of colorectal epithelial cells.

In the intestinal epithelium, TCF4 is the most prominently expressed TCF family member. Gene disruption in the murine germ line has revealed that during embryonic development TCF4 is required to establish the proliferative
5 progenitors of the prospective crypts in the small intestine.

To better understand the contribution of constitutive β-catenin/TCF activity to the colorectal transformation process, the present inventors have
10 undertaken a large-scale analysis of the downstream genetic program activated by β-catenin/TCF in CRC cells.

During this research it was found that inhibition of β-catenin/TCF activity in fully malignant colorectal cancer cells causes these cells to arrest in G1. DNA array
15 analysis revealed the down-regulation of a small set of transcripts. These genes were expressed in polyps, but also in the normal proliferative compartment of colon crypts. The presence of nuclear β-catenin in this compartment was demonstrated, suggesting that WNT signaling is controlling
20 the self-renewing amplification compartment in the adult intestine. In addition, the induction of multiple marker genes of intestinal differentiation upon inhibition of β-catenin/TCF in CRC cells was observed. It was also found that the cell cycle inhibitor p21$^{CIP1/WAF1}$ is an important
25 mediator of this effect. It was concluded that β-catenin/TCF inhibits differentiation and imposes a crypt progenitor-like phenotype on CRC cells.

Moreover, disruption of β-catenin/TCF-activity in CRC cells was shown to restore the physiological program of
30 epithelial differentiation, despite the presence of multiple other mutations present in these cells.

Thus, a group of target genes was identified whose expression is regulated by TCF/β-catenin complexes. In colon carcinoma cells, TCF/β-catenin signalling is deregulated and the resulting inappropriate expression of

5 these target genes is considered to promote carcinogenesis. The transactivation of TCF target genes induced by mutations in APC or β-catenin is believed to represent the primary transforming event in colorectal cancer.

The identification of the target genes of the

10 TCF/β-catenin signalling pathway provides the opportunity to develop therapeutical compounds or therapies that restore or neutralize the inappropriate expression of these genes when TCF/β-catenin signalling is deregulated. By normalizing the expression pattern of one or more of the

15 target genes the drugs can halt or reverse the further development of existing cancer cells, such as colon carcinoma cells, for example by the induction of differentiation of the cancer cells, thus restoring the normal cycle of events.

20 The interference in the inappropriate expression of the target genes can be achieved via the expressed proteins and/or via the transcripts of the genes. These two ways require different active molecules as will be explained herein below.

25 According to a first aspect, the present invention relates to the use of these target genes and/or their expression products for the development of therapeutical compounds, in particular antibodies, small molecules, antisense molecules and/or RNAi molecules, and

30 gene therapies for treating cancers in which TCF/β-catenin/WNT signalling is deregulated, in particular colorectal cancer and melanomas.

This is achieved according to a first embodiment of the invention by characterizing the expression product of the target gene and the production of antibodies against the expressed proteins or peptides derived therefrom and/or

5   of small molecules that bind the expressed protein in a way that inhibits or abrogates its biological function.

According to a second embodiment, the target gene sequence information is used to design antisense molecules, RNAi molecules or gene therapies.

10   A further aspect of the present invention relates to the use of the target genes or their expression products for the development of reagents for diagnosis of cancers in which TCF/β-catenin/WNT signalling is deregulated.

The target genes that were identified according

15   to the invention are the following: CD44, KIT, G protein-coupled receptor 49 (GPR49), Solute Carrier Family 12 member 2 (SLC12A2), Solute Carrier Family 7 member 5, Claudin 1(CLDN1), SSTK serine threonine kinase, FYN oncogene, EPHB2 receptor tyrosine kinase, EPHB3 receptor

20   tyrosine kinase, EPHB4 receptor tyrosine kinase, ETS2, c-Myc, MYB, ID3, POLE3, Bone Morphogenetic Protein 4 (BMP4), Kit ligand (KITLG), GPX2, GNG2, CDCA7, ENC1, the gene identified with Celera ID hCG40185, the gene identified with Celera ID hCG1645335, the gene represented by IMAGE

25   clone 1871074, the gene identified with Celera ID hCG27486, the gene represented by IMAGE clone 294873, the gene represented by IMAGE clone 940994, the gene identified with Celera ID 39573, the gene represented by IMAGE clone 753028, the gene identified with Celera ID hCG37727, the

30   gene identified with Celera ID hCG40978, and the gene identified with Celera ID hCG1811066. Table 1 gives an overview of these target genes. The sequences of these target genes and their expression products are given in the

figures. According to the present invention, target genes which are preferably used comprise a cDNA sequence as shown in Figures 17-24, preferably a sequence, which is at least 90% homologous to the sequences as shown in the Figures 17-

5   24. In addition, the expressed proteins according to the invention preferably comprise a protein sequence as shown in Figures 17-24, preferably a sequence which is at least 90% homologous to the protein sequences shown in Figures 17-24.

10   Based on the above TCF/β-catenin target genes, novel therapeutic compounds and therapies are developed for the treatment of cancer, in particular colorectal cancer and melanomas. Such therapeutic compounds are preferably antibodies, small molecule inhibitors, antisense molecules

15 or RNAi molecules. In addition gene therapies are provided.
          Such gene therapies are based on the generation of dominant-negative (dn) forms of the target genes, which inhibit the function of their wild-type counterparts following their directed expression in a cancer cell.

20 Promoters for use in gene therapy that are specifically activated by TCF/β-catenin to drive specific expression of dominant-negative or suicide genes in cancer cells with active TCF/β-catenin signalling are known from e.g. Lipinski et al. (Mol. Ther. 2001 4:365 - High level β-

25 catenin/TCF dependent transgene expression in secondary colorectal cancer tissue), Chen & McCormick (Cancer Res. 2001 61:4445 - Selective targeting to the hyperactive β-catenin/TCF factor pathway in colon cancer cells), and Fuerer & Iggo (Gene Ther. 2002 9:270 - Adenoviruses with

30 TCF-binding sites in multiple early promoters show enhanced selectivity for tumor cells with constitutive activation of the Wnt signalling pathway).

　　　　In table 1, genes identified as target of TCF/β-
catenin signalling are summarized. Individual target genes
are identified by their official gene symbol and name
(unless otherwise stated) as approved by the Human Gene
5　Nomenclature Committee (HGNC), their gene (hCG), Transcript
(hCT), Protein (hCP), identification code as referred to in
the Celera Discovery System Database
(www.celeradiscoverysystem.com) [or their Genbank mRNA and
protein ID where stated] and their chromosome localization.
10　The genes are classified into broad functional groups
according to their proposed function. Table 1 also shows
the magnitude of down-regulation of expression levels
following inhibition of TCF/β-catenin signalling through
expression of dominant-negative TCF-4 (dnTCF4) in LS174T
15　colon carcinoma cells (ND: not determined).

Table 1

Human Gene Symbol/ Clone ID	Human Gene Name	Celera Gene ID	Celera Transcript ID	Celera Protein ID	Chromo-Some	Group	LS174T dnTCF4 fold down-regulation		confirmed by Northern
							11 hr	23 hr	
CD44	CD44 antigen	hCG1811182	hCT1951772	hCP1753227	11	Membrane (Receptor)	1.8	3.2	Yes
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homologue	hCG22160	hCT13253	hCP39638	4	Membrane (Receptor)	ND	ND	Yes
GPR49	G protein-coupled receptor 49	hCG23766	hCT14878	hCP42243	12	Membrane (Receptor)	2.3	3.6	Yes
SLC12A2	Solute carrier family 12, member 2	hCG27034	hCT18167	hCP44322	5	Membrane	1.8	4.2	Yes
SLC7A5	Solute carrier family 7, member 5	hCG1789357	hCT1828603	hCP1742844	16	Membrane	1.5	3.6	Yes
CLDN1	Claudin 1	hCG17574	hCT8625	hCP33694	3	Membrane	1.2	3.0	ND
SSTK	Serine/threo-nine protein kinase	hCG1640387	hCT1640514	hCP1611903	19	Kinase	1.1	5.6	ND
FYN	Fyn oncogene related to SRC, FGR, YES	hCG34806	hCT26018	hCP47031	6	Kinase	1.3	2.7	ND
EPHB2	Eph-related receptor tyrosine kinase B2	hCG1812037	hCT1955735	hCP1765674	1	Kinase	1.8	3.0	Yes
EPHB3	Eph-related receptor tyrosine kinase B3	hCG16839	hCT7881	hCP33573	3	Kinase	2.9	6.6	Yes

Gene	Description	hCG	hCT	hCP		Class			
EPHB4	Eph-related receptor tyrosine kinase B4	hCG20448	hCT11528	hCP38155	7	Kinase	ND	ND	Yes
ETS2	v-ets erythroblastosis virus E26 oncogene homologue 2 (avian)	hCG401219	hCT401223	hCP801143	21	Transcription	1.5	3.4	Yes
c-Myc	v-myc myelocytomatosis viral oncogene homologue (avian)	hCG15917	hCT6947	hCP33451	8	Transcription	1.5	3.3	Yes
MYB	v-myb myeloblastosis viral oncogene homologue (avian)	hCG32380	hCT23568	hCP46310	6	Transcription	1.6	3.3	Yes
ID3	Inhibitor of DNA binding 3, dominant negative helix-loop-helix protein	hCG1739237	hCT1955731/hCT1955730/hCT1955731/hCT1777319	hCP1765651/hCP1765654/hCP1713380	1	Transcription	2.2	2.9	Yes
POLE3	Polymerase (DNA directed), epsilon 3 (p17 subunit)	hCG29189	hCT20349	hCP44379	9	Transcription	ND	ND	ND
BMP4	Bone morphogenetic protein 4	hCG20967	hCT1955929/hCT12051	hCP1765419/hCP40076	14	Secreted Factor	3.0	3.6	Yes
KITLG	KIT Ligand	hCG26603	hCT18442/hCT18443	hCP42895/hCP42896	12	Secreted Factor/membrane	ND	ND	Yes
GPX2	Glutathiane peroxidase 2 (gastrointestinal)	hCG40439	hCT31698	hCP51533	14	enzyme	1.2	3.5	ND
GNG2	Guanine nucleotide binding protein (G protein), gamma 2	hCG22671	hCT13769	hCP40131	14	Signaling	0.9	3.2	Yes

CDCA7	Cell-division cycle associated 7	hCG16803	hCT7844	hCP34351	2	Unknown (nuclear)	ND	ND	Yes
ENC1	Ectodermal-neural cortex	hCG37104	hCT1957908/hCT40185	hCP48121/hCP176097	5	Unknown	4.1	2.6	Yes
STXBP6	Syntaxin binding protein 6 (amisyn)	hCG40185	hCT1813210/hCT31439	hCP1768217/hCP50010	14	Unknown	ND	ND	ND
Image ID 1048671	Unknown	hCG1645335	hCT1645462	hCP1634454	X	Unknown	ND	ND	Yes
Image ID 1871074	Unknown	Unknown	Unknown	Unknown	2	Unknown	ND	ND	Yes
Image ID 376697	LOC285705 (PROVISIONAL)	hCG27486	hCT18626	hCP43057	5	Unknown	ND	ND	Yes
Image ID 294873	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	ND	ND	ND
Image ID 940994	Unknown	Unknown	Unknown	Unknown	11	Unknown	ND	ND	Yes
Image ID 742837	LOC286459 (PROVISIONAL)	hCG39573	hCT30826	hCP49367	X	Unknown	ND	ND	ND
Image ID 753028	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	ND	ND	ND
TM4SF9	Tetraspan 5	hCG37727	hCT28961	hCP48431	4	Membrane	ND	ND	ND
Image ID 294133	FLJ12747 (PROVISIONAL)	hCG40978	hCT32246	hCP50788	22	Unknown	ND	ND	ND
Image ID 489594	FLJ11565 (PROVISIONAL)	hCG1811066	hCT1951376	hCP1751927	X	Unknown	ND	ND	ND

A method for the development of the therapeutic compounds according to the invention comprises the steps of:

a) identification of genes regulated by TCF/β-catenin in colon carcinoma cells, in particular by using microarray technologies;

b) validation of one or more of the identified genes as potential target gene(s) for the therapeutic compound by one or more of the following methods:

- confirmation of the identified gene by Northern Blot analysis in colon carcinoma cell-lines;

- determination of the expression profile of the identified gene in human colorectal tumors and normal tissue;

- determination of the functional importance of the identified target genes for colorectal cancer;

c) production of the expression product of the target gene; and

d) use of the expression product of the target gene for the production or design of a therapeutic compound.

Once the target gene is validated and the expression product of the gene (the expressed protein) is produced there are various ways for developing a therapeutic compound for treating colorectal cancer.

In colorectal carcinoma cells the TCF/β-catenin regulated target genes identified according to the invention are over-expressed upon constitutive TCF/β-catenin activity. The compounds of the invention should thus neutralize the biological activity of the proteins expressed by the target gene in order to reverse the

carcinoma phenotype.

A known way of neutralizing proteins is by means
of antibodies. The invention according to a first aspect
thereof thus relates to antibodies directed against the
5   expression products of the target genes listed in Table 1
for use in the treatment of colorectal cancer. The
production and evaluation of antibodies and their
derivatives, such as scFv, Fab, chimeric, bifunctional and
other antibody-derived molecules are well within the reach
10   of the skilled person. Therapeutic antibodies are in
particular useful against target gene expression products
located on the cellular membrane.

A second aspect of the invention relates to so-
called "small molecules" interfering with the biological
15   activity of the protein expressed by the target gene for
use in the treatment of colorectal cancer. Small molecules
are usually chemical entities that are developed on the
basis of structure-function analysis of the protein with
which they should interfere. Such analysis may involve
20   determination of the crystal structure of the target
protein. Based on the information thus obtained libraries
of compounds can be screened or compounds may be designed
and synthesized using medicinal and/or combinatorial
chemistry. Alternatively, high throughput screening can be
25   used to generate useful drug lead compounds as well. After
identification of a lead compound, this compound is
screened for inhibition of target protein function
using in vitro and/or cell-based assays. After optimization
of the lead compound with respect to its structure,
30   toxicity profile and inhibition capability, its efficacy as
colon cancer therapeutic is tested in vivo using animal
models (e.g. Xenograft, APC^{min} mouse).

According to a third aspect of the invention

antisense molecules are provided. Antisense drugs are
complementary strands of small segments of mRNA. To create
antisense drugs, nucleotides are linked together in short
chains called oligonucleotides. Each antisense drug binds
5   to a specific sequence of nucleotides in its mRNA target to
inhibit production of the protein encoded by the target
mRNA. By acting at this earlier stage in the disease-
causing process to prevent the production of a disease-
causing protein, antisense drugs have the potential to
10   provide greater therapeutic benefit than traditional drugs
which do not act until the disease-causing protein has
already been produced. The invention relates to antisense
molecules directed against the target genes listed in Table
1.
15   A further aspect of the invention relates to RNA
interference (RNAi) molecules. RNAi refers to the   .
introduction of homologous double stranded RNA to
specifically target a gene's product, resulting in a null
or hypomorphic phenotype. RNA interference (RNAi) requires
20   an initiation step and an effector step. In the first step,
input double-stranded (ds) RNA is processed into 21-23-
nucleotide "guide sequences". These may be single- or
double-stranded. The guide RNAs are incorporated into a
nuclease complex, called the RNA-induced silencing complex
25   (RISC), which acts in the second effector step to destroy
mRNAs that are recognized by the guide RNAs through base-
pairing interactions. RNAi molecules are thus double
stranded RNAs (dsRNAs) that are very potent in silencing
the expression of the target gene. Potentially, a single
30   dsRNA molecule could mark hundreds of mRNAs for
destruction.
              The invention relates further to gene therapy, in
which the target genes are used for the design of dominant-

negative genes which inhibit the function of the
corresponding target gene following their specific
expression in a cancer cell. Alternatively, RNAi approaches
can be used gene therapeutically, for example by
5   introducing a dsRNA producing sequence into a cancer cell.

The invention further relates to pharmaceutical
compositions for treating cancers in which TCF/β-catenin
signalling is deregulated, in particular colorectal cancer
and melanomas, said compositions comprising a suitable
10 excipient, carrier and/or diluent, and one or more
inhibitors of the proteins expressed by the TCF/β-catenin
target genes, or inhibitors of the mRNAs of the target
genes.

The invention also provides diagnostic methods
15 for diagnosing cancer, in particular colorectal cancer and
melanomas, comprising histological examination of tissue
specimens, using specific antibodies directed against
TCF/β-catenin target gene products and/or in situ
hybridisation analysis of TCF/β-catenin target gene
20 expression using specific RNA probes directed against
TCF/β-catenin target genes.

The present invention will be further illustrated
by the following, non-limiting, Examples. In the Examples
reference is made to the following figures:
25   **Figure 1.** TCF/β-catenin driven transactivation is
abrogated upon induction of dominant-negative TCF (dnTCF).

(A) Inducible dnTCF4 expression in the CRC line
Ls174T. Cells were stained for dnTCF4, 24 hours after
induction with doxycycline. DnTCF4 is highly expressed in
30 the nucleus.

(B) dnTCF4 protein is induced as early as 4 hours
after induction with doxycycline as analysed by western

blot.

(C) Both dnTCF1 and dnTCF4 abrogate β-catenin/TCF
driven ·transcription in the β-catenin-mutant Ls174T as well
as in the APC-mutant DLD1 cells. Activity of the TCF-
5   reporter TOPFlash (purple bars) and control FOPFlash (green
bars) after 24 hours with or without doxycycline (dox) is
shown. Parental cell lines expressing the Tet-repressor
alone are used as controls.

Figure 2. Northern blot analysis of genes
10 regulated by β-catenin/TCF activity.

(A)   Representative examples of several
        Tcf/β—catenin target genes in Ls174T and
        DLD-1 colon carcinoma cells. The indicated
        mRNAs were down-regulated following 24 hours
15   of doxycycline-induced expression of
        dominant-negative Tcf (dnTCF). The bottom
        panel shows the 28S ribosomal RNA as a
        loading control. (Dox: Doxycycline; CON: RNA
        from control cells lacking dominant-negative
20   Tcf expression).

(B)   Additional examples of genes whose
        expression is dependent upon active TCF/β-
        catenin signalling in LS174T colon cancer
        cell-lines. The indicated mRNAs were down-
25   regulated following 24 hours of doxycycline
        induced expression of dominant-negative TCF.

Figure 3. Expression of dnTCF induces G1-cell
cycle arrest in colon carcinoma cells.

(A) Ls174T and DLD1 show a dramatic reduction in
30 S phase cells upon dnTCF expression. The scatter profile
of cells in G1 (blue), S (green) and G2/M (red) after 20
hours with or without doxycycline is shown. Numbers refer

to the percentage of cells in S phase for each cell line analyzed. The results are representative of several independent experiments.

(B) Proliferation was halted in Ls174T and DLD1 transfectants. This was visualized by crystal violet staining of cell cultures after 5 days of dnTCF expression.

Figure 4. (A-C) The expression of nuclear β-catenin (A, black arrowheads) perfectly correlates with that of EPHB2 tyrosine kinase receptor (B, black arrowheads) in aberrant crypt foci (ACF). Stainings were performed on serial sections of early human lesions. The dashed lines delimit the same ACF in both stainings. EPHB2 is expressed at the bottom of the crypts (C, white arrowheads).

Figure 5. Model for the role of β-catenin/TCF in the early stages of intestinal tumorigenesis.

(A) Schematic representation of a colon crypt and proposed model for polyp formation. At the bottom third of the crypt, the progenitor proliferating cells accumulate nuclear β-catenin. Consequently they express β-catenin/TCF target genes. An uncharacterised source of WNT factors likely resides in the mesenchymal cells surrounding the bottom of the crypt, depicted in red. As the cells reach the mid-crypt region, β-catenin/TCF activity is downregulated and this results in cell cycle arrest and differentiation. Cells undergoing mutation in *APC* or β-catenin become independent of the physiological signals controlling β-catenin/TCF activity. As a consequence, they continue to behave as crypt progenitor cells in the surface epithelium giving rise to ACFs.

(B) CD44, a β-catenin/TCF target, exemplifies this model. It is expressed in the normal proliferative

compartment at the bottom of the crypts (white arrowheads)
and also in the early lesions arising at the surface
epithelium (black arrowheads).

Figure 6. Expression of β-catenin/TCF target
5   genes in normal colon and colorectal polyps.

(A-F) Immunohistochemical analysis of the
expression of Bmp4 (A and B), cMyb, (C and D) and Enc1 (E
and F) in normal mouse colon (A, C and E) or colorectal
polyps from *min* mice (B, D and F). Target genes are highly
10   expressed at the bottom of the normal crypts (white
arrowheads) and in colorectal polyps arising at the surface
epithelium (black arrowheads).

Figure 7. Expression of EPHB3 and EPHB4 in normal
colon and colorectal polyps. (D,F,G) EPHB3 and EPHB4 are
15   expressed in the crypts of normal mouse colon (White
arrowheads, D,E respectively) and are over-expressed in
intestinal polyps of APCmin mice (Black arrowheads, F,G
respectively). EPHB3 expression is absent in TCF-4 knock-
out mice deficient in TCF-4/β-catenin signalling in the
20   small intestine(H).

Figure 8. (A) Semi-quantitative RT-PCR analysis
of selected TCF target gene expression levels in human
cancer cell-lines. (B) Summary table of estimated TCF
target gene expression levels in various cancer cell-lines
25   (- undetectable expression, + low level expression, ++
moderate expression, +++ high-level expression, +++ very
high level expression, ND not determined, L 100bp DNA
ladder).

Figure 9. Quantification of selected TCF target
30   gene expression levels in human cancer cell-lines.

(A) Endogenous GPX2 mRNA levels in a panel of 20
human cancer cell lines of varying origin. Q-PCR analysis

of endogenous GPX2 expression levels reveals high-level
expression in cancer cell-lines `of varying origin, with
particularly high levels evident in colon carcinoma cell-
lines LS174T, HT29, SW480, the lung cancer cell-line A549

5   and the breast cancer cell-line MDAMB351. Numbers above the
figure correspond to very high GPX2 mRNA values not shown
to scale.

(B) Endogenous EPHB2 mRNA levels in a panel of 20
human cancer cell lines of varying origin. Q-PCR analysis

10   of endogenous GPX2 expression levels reveals high-level
expression in cancer cell-lines of varying origin, with
particularly high levels evident in colon carcinoma cell-
lines LS174T, DLD-1 and the prostate cancer cell-line PC3.
Numbers above the figure correspond to very high EPHB2 mRNA

15   values not shown to scale.

(C) Endogenous RGMR mRNA levels in a panel of 20
human cancer cell lines of varying origin. Q-PCR analysis
of endogenous RGMR expression levels reveals high-level
expression in selected colon, lung and prostate cancer

20   cell-lines, with highest levels evident in the colon cancer
cell-line LS174T.

(D) Endogenous TSpan5 mRNA levels in a panel of
20 human cancer cell lines of varying origin. Q-PCR
analysis of endogenous Tspan5 expression levels reveals

25   high-level expression in selected colon, lung and prostate
cancer cell-lines, with highest levels evident in the colon
cancer cell-line HCT116.

Figure 10. (A) Schematic representation of the
CD44 gene. Open boxes indicate exons that can be

30   alternatively spliced. TM: transmembrane region.

(B, C) Schematic representation of the CD44
protein with localizations of the epitopes that are
recognized by the anti-human monoclonal antibodies VFF18

and Hermes-3 and the anti-mouse antibodies PGP-1, 10D, and 9A4. v1 to v10, domains encoded by variant exons.

Figure 11. Exon/Intron organization of human CD44. Exons encode the constant part of the extracellular

5   domain. Exons 6-15 correspond to variant exons 1-10 respectively and encode for the extracellular domain (Variant exon 1 is not expressed in humans). Exons 16 and 17 are constant exons, which together with part of exon 5 encode the membrane proximal region of the extracellular

10   domain. Exon 18 encodes the hydrophobic transmembrane region and exons 19 and 20 encode the cytoplasmic domain. Exons 19 and 20 are also subject to alternative splicing, generating either long or short cytoplasmic domains.

Figure 12. CD44 expression requires TCF/β-catenin

15   signalling.

(A) Inhibition of TCF/β-catenin signalling using dnTCF results in loss of CD44 expression on the cell-surface of colon cancer cells (Pan-CD44Mab used).

(B) CD44 expression is lost in intestinal crypts

20   following deletion of TCF-4 in mice.

Figure 13. CD44 is over-expressed in early colorectal polyps in comparison to normal colon. Stainings were generated using a Pan-CD44 antibody.

Figure 14. Schematic representation of GPR49

25   which belongs to the G protein-coupled receptor (GPCR) superfamily with a large seven-transmembrane (TM).

Figure 15. Lineup of RGM and RGMR Protein Sequences.

Figure 16. Schematic representation of Tspan5,

30   which comprises 4 transmembrane domains and two large extracellular loops.

Figure 17. cDNA and protein sequence of CD44: A. cDNA Sequence (hCT1951772/NM_000610), B. Protein

sequence (hCP1753227/ NP_000601).

Figure 18. (A) cDNA sequence: X55150 (alternative splice variant CD44E), (B). Protein sequence: S13530(alternative splice variant CD44E).

5   Figure 19. CD44 alternative splice variants containing any combination of the following variable exons: A. Variable exon 2 (L05411), B. Variable exon 3 (L05412), C. Variable exon 4 (L05413), D. Variable exon 5 (L05414), E. Variable exon 6 (L05415), F. Variable exon 7 (L05416), 10 G. Variable exon 8 (L05417), H. Variable exon 9 (L05418), I.Variable exon 10 (L05419).

Figure 20. cDNA and protein sequence of GPR49: A. cDNA Sequence: hCT14878, B. Protein Sequence: hCP42243

15   Figure 21. cDNA and protein sequence of EPHB4: A. cDNA Sequence: hCT11528/ NM_004444, B. Protein Sequence, hCP38155/ NP_004435.

Figure 22. cDNA and protein sequence of GPX2: A. cDNA sequence: X68314, B. Protein Sequence: 20 CAA48394.

Figure 23. cDNA and protein sequence of hCG27486 (RMGR): A. cDNA sequence: hCT18626, B. Protein Sequence: hCP43057.

Figure 24. cDNA and protein sequence of Tspan5: 25 A. cDNA sequence: NM_005723, B. Protein sequence: NP_005714.

EXAMPLES

EXAMPLE 1

Identification of target genes

MATERIAL AND METHODS

Cell culture and transfections

Cells were grown in RPMI supplemented with 10% FCS and antibiotics. T-REx system (Invitrogen) was used according to manufacturer's instructions to generate inducible dnTCF or p21$^{CIP1/WAF1}$ inducible CRC cell lines. In short, 10^7 cells were transfected by electroporation with 20 µg FspI linearized pcDNA$_6$TR . After 3 weeks of selection, blasticidin (10 µg/ml) resistant colonies were expanded and transfected with pcDNA$_4$TO-Luciferase. From each cell line, two clones showing the strongest induction were chosen. These were subsequently transfected with 20 µg PvuI linearized dnTCF1 or dnTCF4 in pCDNA$_4$TO. After selection on Zeocin (500µg/ml), resistant colonies were tested for dnTCF induction by immunocytochemical staining after addition of doxycycline and selected for further studies.

Cell cycle analysis

3×10^6 (Ls174T) or 10^6 (DLD1) cells were seeded in 9 cm dishes and doxycycline was added (1µg/ml). After 20 hrs, BrdU (Roche) was added for 20 min. Cells were then collected and fixed in ethanol 70%. Nuclei were isolated, incubated with α-BrdU-FITC (BD) and cell cycle profiles were determined by FACS analysis. Crystal violet staining on methanol fixed cells was performed on cells after 5 days in culture with or without the addition of doxycycline.

RNA isolation and northern analysis

RNA was isolated using Trizol (Gibco). Northern blots were performed according to standard procedures .
5   Probes were obtained by appropriate restriction enzyme digestion of corresponding IMAGE clones (IMAGE consortium) spotted in the array.

Immunohistochemistry

10   The antibodies used in this study were obtained from the following sources: EPHB2, EPHB3 and EPHB4 from R&D systems; BMP4 from Novacastra; ENC1 from Pharmingen; MYB from Santa Cruz Biotechnology; p21$^{CIP1/WAF1}$ from Pharmingen; carbonic anhydrase II from Rockland; β-catenin from
15   Transduction Laboratories; TCF1 and TCF4 antibodies were described elsewhere. Immunostainings were performed according to standard procedures. Briefly, sections were pretreated with peroxidase blocking buffer (100 mM Na-phosphate pH 5.8, 30 mM NaN$_3$, 0.2% H$_2$O$_2$) for 20 minutes at
20   room temperature after dewaxing and hydration. Antigen retrieval was performed by boiling samples in 10 mM Na-citrate buffer pH 6.0, for 20 minutes. For β-catenin stainings, samples were boiled for 45 minutes in 40 mM Tris pH 8.0 containing 1 mM EDTA. Incubation of antibodies was
25   performed in 1% BSA in PBS 1 hour at room temperature. In all cases, the Envision+ kit (DAKO) was used as a secondary reagent. Stainings were developed using DAB (brown precipitate). Slides were then counterstained with hematoxylin and mounted.

Probe preparation and microarray procedures

mRNA was extracted from cells using the Fasttrack 2.0 procedure (Invitrogen Inc.) following the

5 manufacturer's directions. Fluorescent labeled cDNA was prepared from 1μg of polyA mRNA by oligo dT-primed polymerization using Superscript II reverse transcriptase in the presence of either Cy3- or Cy5- labeled dCTP as described (website:

10 http://cmgm.stanford.edu/pbrown/protocols.html). The appropriate Cy3- and Cy5- labeled probes were pooled and hybridized to microarrays in a volume of 25 μl under a 22x14 mm glass coverslip for 16 hr. at 65°C and washed at a stringency of 0.2XSSC. The microarray contains 24,000 DNA

15 spots representing approximately 10,000 known full-length cDNAs and 14,000 ESTs of clones made available by Research Genetics, which are listed in the supplementary information.

Fluorescent images of hybridized microarrays were

20 obtained using a genepix 4000 microarray scanner (Axon Instruments, Inc). Images were analyzed with scanalyze (M.Eisen; http://www.microarrays.org/software) or with genepix 3.0. Fluorescence ratios were stored in a custom database. Fluorescent ratios were calibrated independently

25 for each array by applying a single scaling factor to all fluorescent ratios from each array; this scaling factor was computed so that the median fluorescence ratio of the measured spots on each array was 1.0. Genes represented by good-quality spots for which the fluorescent intensity in

30 each channel was greater than 1.5 times the local background were selected.

RESULTS

**Generation and characterization of inducible
dnTCF cell lines**

5   To determine the role of β-catenin/TCF complexes
in established CRC cells, cell lines were constructed
carrying doxycycline-inducible expression plasmids encoding
N-terminally truncated versions of TCF factors. Such
dominant-negative TCF (dnTCF) proteins do not bind β-
10   catenin and therefore act as potent inhibitors of the
endogenous β-catenin/TCF complexes present in CRC. As the
recipient cell line, the CRC cell line Ls174T, which
expresses mutant β-catenin protein, yet is diploid and
carries wild-type alleles of *p53* and *APC* was initially
15   chosen. Multiple clones were isolated and tested for
inducibility of dnTCF4 expression.

          Strong nuclear dnTCF4 staining was observed after
doxycycline (Dox) induction of positive transfectants
(Figure 1A). Accumulation of the induced protein could be
20   detected as early as 4 hours after the addition of
doxycycline (Figure 1B). CRC cell lines such as Ls174T that
carry WNT pathway mutations constitutively activate TCF
reporters (pTOPFlash). Several clones were selected in
which the inducible expression of dnTCF4 completely
25   abrogated this constitutive pTOPFlash activity (Figure 1C).
Induction of dnTCF4 in such clones imposed a robust cell
cycle block (see below), but did not result in the onset of
apoptosis.

30   The genetic program under the transcriptional
control of β-catenin/TCF activity in CRC cells

          The spectrum of target genes controlled by β-
catenin/TCF in CRC cells was expected to hold the key to

understanding the primary transformation of intestinal
cells. By DNA array analysis, it was determined which genes
were specifically affected in their expression upon the
induction of dnTCF4. mRNA was isolated at 11 hours and 23

5   hours after the initiation of the experiment with or
without the addition of doxycycline. cDNA prepared from the
uninduced samples was labeled with Cy3, while the induced
cDNA samples were labeled with Cy5. At each time point, the
uninduced and induced cDNA samples were mixed and

10   hybridized in duplicate to a DNA array consisting of
approximately 24,000 cDNA spots representing known genes or
EST clusters. Fluorescent images were analysed as detailed
in the experimental procedures.

        A single criterium was applied to the array data

15   set: i.e. a decrease of at least 2.5 fold in both
measurements at the 23 hour time point. This defined a
small set of 35 entries that were downregulated when β-
catenin/TCF activity was abrogated in Ls174T cells
expressing dnTCF-4 (listed in Table 1).

20   For a number of downstream genes defined in the
Ls174T cells expressing dnTCF4, Northern blot analysis was
performed before and after induction of dnTCF4. This
invariably confirmed the DNA array data (Table I and Figure
2A/2B). The down-regulation of the reported TCF4 target

25   gene c-MYC did not meet the 2.5-fold selection criterion,
decreasing by an average 1.8-fold. However, its relatively
modest, but consistent down-regulation was confirmed by
northern blot (Figure 2A).

        To further investigate the effects of β-

30   catenin/TCF inhibition, Ls174T cells expressing dnTCF1, the
natural dominant-negative isoform of TCF1 expressed in the
intestinal epithelium, were constructed. Likewise, DLD-1

cells, a CRC cell line with wild type β-catenin but mutated *APC* and *p53*, was engineered to express inducible dnTCF1 or dnTCF4.

5   Again, doxycycline-induced expression of dnTCFs in all cell lines resulted in the abrogation of pTOPflash activity (Figure 1C) and cell cycle arrest (see below). Northern blot analysis was performed in the above dnTCF expressing cell lines. Almost invariably, the target genes listed in Table I were also strongly downregulated by

10   dnTCF1 in Ls174T (Figure 2A). In addition, the DLD-1 cells showed a similar pattern of target gene expression upon inhibition of β-catenin/TCF activity by dnTCF1 or dnTCF4 (Figure 2A).

15   **Inhibition of β-catenin/TCF activity leads to cell cycle arrest and differentiation in CRC cells.**

The induction of dnTCF4 or dnTCF1 in both Ls174T and DLD1 cell lines had a dramatic effect on cell cycling. Within 20 hours, a robust G1 arrest was induced (Figure

20   3A). Accordingly, cell proliferation was halted upon doxycycline induction of dnTCFs as visualized by crystal violet staining of cell cultures induced for 5 days (Figure 3B).

25   **The genetic program controlled by β-catenin/TCF in CRC cells is physiologically active in colonic epithelium**

In order to validate the β-catenin/TCF target genes described in this example, immunohistochemical

30   analyses of those entries for which antibodies were available were performed on early intestinal neoplastic lesions. In Figure 4, a representative example of this

analysis is shown. As expected, a strict correlation
between the accumulation of nuclear β-catenin (Figure 4A)
and the expression of EPHB2 (Figure 4B) was observed in
early colorectal lesions. Many other downregulated genes

5   listed in Table I were overexpressed in early intestinal
polyps from *Min* mice or in aberrant crypt foci (ACF) from
FAP patients (Figure 5, Figure 6, Figure 7).

        More strikingly, it was found that EPHB2 was not
only expressed in polyps, but also in cells within the

10 proliferative compartment at the bottom of normal colon
crypts (Figure 4C). This pattern was invariably confirmed
for all target genes tested by immunohistochemistry. These
included *MYB, BMP4, ENC1,* (Figure 6), *EPHB3* (Figure 7), and
CD44 (Figure 5B).

15   Thus, the observed gene expression changes in CRC
cells recapitulated the physiological differentiation of
crypt progenitor cells during their migration towards the
luminal surface of the intestine.

20 *DISCUSSION*

        The data presented here provide a view of the
genetic program driven by β-catenin/TCF activity in CRC
cells. The expression of a surprisingly limited set of

25 genes is dependent on the presence of active β-catenin/TCF
complexes.

        A hallmark of cancer is deregulated
proliferation. Abrogation of β-catenin/TCF activity in all
CRC cell lines tested induced a robust arrest in the G1

30 phase of the cell cycle, demonstrating that the activity of
the β-catenin/TCF complex represents the major force
driving cell proliferation in intestinal cells.

β-Catenin and TCF modulate cell cycle control by activating genes that promote cell cycling (e.g. c-myc), but also by repressing cell cycle inhibitors (p21$^{CIP1/WAF1}$; results not shown). β-catenin/TCF represents the main

5   upstream regulator of the cell cycle machinery in epithelial intestinal cells.

In conclusion, the above observations demonstrate that TCF constitutes the dominant switch between the proliferating progenitor and the differentiated intestinal

10 cell. This is recapitulated in the CRC cells used in this study, despite the presence of multiple additional mutations in these cells. This example validates that the genetic program controlled by TCF/β-catenin signalling can be used as the basis for the development of a therapeutic

15 strategy to revert the transformed phenotype in colorectal cancer.

EXAMPLE 2
Development of drugs for the treatment of colorectal

20 cancer on the basis of the target genes of Example 1

Identification and validation of the target genes
Example 1 demonstrates the identification of

25 target genes represented on cDNA/oligonucleotide microarrays which are regulated by TCF/β-catenin transcription factor complexes. Subsequently, the regulated expression of target genes in colon cancer cell-lines is confirmed via Northern blot analysis using gene specific

30 probes as described in Example 1.

In order to confirm that the expression of the target genes that were found in Example 1 are linked to the TCF/β-catenin complex, target gene expression is also

evaluated in tissues known to have active TCF/β-catenin
complexes (for example, intestinal epithelium and
colorectal polyps) using gene-specific antibodies, <u>in situ</u>
hybridization with gene-specific probes and/or RT-PCR with
5   gene-specific primers.

After that, the expression profile of the target
gene in human/mouse cell-lines and tissues is determined
via Northern blot analysis and/or RT-PCR. This is done
because ubiquitous expression of the target gene may be
10 indicative of possible side-effects of therapeutics
designed to block the target gene's function <u>in vivo</u>.

Obtaining the complete gene

The identification of the target genes on a
15 microarray does not identify the complete gene. The next
step in the development of a therapeutic compound is
therefore the generation of full-length clones for the EST
sequences which are shown to be regulated via TCF/β-catenin
in the colon carcinoma cell-lines. This is achieved by
20 searching databases for full-length EST clones and/or
techniques such as RT-PCR, RACE and hybridization screening
of cDNA libraries.

Identification of binding sites within the target
25 genes

Putative TCF binding sites [(A/T)(A/T)CAA(A/T)GG]
within target gene promoters are identified according to
Van de Wetering et al. (Identification and cloning of TCF-
1, a T-lymphocyte-specific transcription factor containing
30 a sequence-specific HMG box; EMBO J. 11: 3039-3044, 1991).
Enhancers are identified using web-based prediction
programs such as Genomatix
(<u>www.genomatix.gsf.de/promoterinspector</u>). This provides an

indication that a gene is regulated via direct binding of
TCF/β-catenin complexes. However, many binding sites will
not be identified due to the vast tracts of genomic DNA
containing the target gene which may harbor distant
5   enhancers. Testing of the functionality of putative TCF
binding sites in target genes is then performed via
mutational analysis. Promoter regions of target genes
containing the original or mutated putative TCF binding
sites are cloned upstream of TK-Luciferase reporter genes
10 cassettes and analysed for their ability to drive
expression of the reporter gene in the presence of TCF/β-
catenin complexes in cultured cell-lines, such as described
in Tetsu and McCormick, (1999) (β-catenin regulates
expression of cyclin D1 in colon carcinoma cells. Nature
15 398: 422-426)). A correlation between mutation of a TCF
binding site and loss of reporter gene expression indicates
that direct binding of TCF/β-catenin is contributing to
target gene expression.
Determination of the ability of ectopic target gene
20 expression to overcome defects in the growth of colon
carcinoma cells caused by blocking TCF/β-catenin signalling
is performed as described in Example 1 to establish whether
expression of a single target gene is sufficient to
overcome the block in cell-cycle and differentiation of the
25 colon cancer cell.

        Confirmation of involvement of target genes in
colon cancer
        Subsequently it is important to establish the
30 contribution of specific target genes to colon cancer. The
techniques used to do this are dominant-negative
approaches, i.e. expression of target genes carrying
deletions/mutations which suppress the function of their

endogenous counterparts in colon cancer cell-lines;
antisense/RNAi approaches, i.e. introduction of double-
stranded RNA oligonucleotides designed to block expression
of a specific target gene into colon cancer cell-lines, as
5   described by Elbashir et al. (Duplexes of 21-nucleotide
RNAs mediate RNA interference in cultured mammalian cells.
Nature 411: 494-498, 2001) and genetic approaches, such as
generation of mice deficient for target gene expression in
intestinal tissues using a combination of standard loxP
10   knockout technologies and intestine-specific Cre mouse
strains, or generation of transgenic mice expressing
dominant-negative target genes. These mice strains are
crossed with APC^{min} mice to determine whether loss of target
gene function in vivo has any adverse effect on colorectal
15   polyp formation (as for example described by Oshima et al.
(Suppression of Intestinal Polyposis in APC^{D716} Knockout
mice by inhibition of cyclooxygenase 2 (COX-2) Cell 87:
803-809, 1996).

Using these approaches it is determined whether
20   loss of function of a specific target gene has any adverse
effect on colon cancer cell-lines and/or on polyp formation
in vivo and thus insight is gained into whether
therapeutics designed to specifically inhibit the function
of these target genes are likely to be effective in
25   combating colon cancer in humans.

Furthermore, the genetic programs affected by
inhibition of target gene function in colon carcinoma cells
are evaluated using microarrays. Thus, the function of the
target gene in colon carcinoma cells is established and
30   valuable information regarding the possible side-effects
that inhibition of this gene function may have on genetic
programs required for normal cell function is provided. By
definition, many of the validated TCF/β-catenin target

genes will be more highly expressed on colon carcinoma tissues than healthy tissues and some encode cancer-specific proteins, making these excellent targets against which to develop colon cancer therapeutics.

5

Identification or development of candidate compounds

Antibodies

Validated target genes which express membrane-
10 bound proteins are then selected as targets for conventional antibody-based therapies, for example according to Schwartzberg (Clinical experience with edrecolomoab: a monoclonal antibody therapy for colorectal carcinoma. Crit. Rev. Oncol. Hematol. 40: 17-24, 2001).

15

Small molecules

Validated intracellular and membrane-expressed target proteins are furthermore selected as targets for developing small molecule compound-based therapies. For
20 this their crystal structures are determined, either from published information available from web-based databases (NCBI) or using protein production and crystallization facilities. Structure analysis is performed with the computer programs SPOCK (Christopher J. (1998). SPOCK, The
25 structural properties observation and calculation kit, 1998), GRASP (Nicholls et al., (1991) Structure, Function and Genetics 11:281-283) and SWISS PDB Viewer (Guez and Peitsch (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling) or others.
30 In addition to the rational development of novel small molecules, high capacity screening of existing small molecule compound libraries generate lead compounds which become inhibitors of validated target proteins encoding

enzymes such as protein kinases. Highly active inhibitors
are co-crystallized with the enzyme and computer programs
such as GOLD (Distributed via Cambridge Crystallographic
Data Centre; Jones et al., (1995) J. Mol. Biol 245: 43-53)
5   and CERIUS2/LUDI (Bohm (1992) The computer program Ludi: A
new method for the de novo design of enzyme inhibitors. J.
Comp. Aided Molec. Design 6:61-78) are used for structure-
based design of improved inhibitors.

10   Antisense molecules
            These can be either antisense RNA or antisense
oligodeoxynucleotides (antisense ODNs), and can be prepared
synthetically or by means of recombinant DNA techniques.
Both methods are well within the reach of the person
15 skilled in the art. ODNs are smaller than complete
antisense RNAs and have therefore the advantage that they
can more easily enter the target cell. In order to avoid
their digestion by DNAse, ODNs, but also antisense RNAs,
are chemically modified. For targeting to the desired
20 target cells the molecules are linked to ligands of
receptors found on the target cells or to antibodies
directed against molecules on the surface of the target
cells.

25   RNAi molecules
            Double-stranded RNA corresponding to a particular
gene is a powerful suppressant of that gene. The ability of
dsRNA to suppress the expression of a gene corresponding to
its own sequence is also called post-transcriptional gene
30 silencing or PTGS. The only RNA molecules normally found in
the cytoplasm of a cell are molecules of single-stranded
mRNA. If the cell finds molecules of double-stranded RNA,
dsRNA, it uses an enzyme to cut them into fragments

containing 21-25 base pairs (about 2 turns of a double
helix). The two strands of each fragment then separate
enough to expose the antisense strand so that it can bind
to the complementary sense sequence on a molecule of mRNA.

5   This triggers cutting the mRNA in that region thus
destroying its ability to be translated into a polypeptide.
Introducing dsRNA corresponding to a particular gene will
knock out the cell's own expression of that gene. This can
be done in particular tissues at a chosen time. A possible

10   disadvantage of simply introducing dsRNA fragments into a
cell is that gene expression is only temporarily reduced.
However, introducing into the cells a DNA vector that can
continuously synthesize a dsRNA corresponding to the gene
to be suppressed can provide a more permanent solution.

15   RNAi molecules are prepared by methods well known to the
person skilled in the art.

Other compounds

        To predict the location of critical contact sites

20   for cofactors, ligands or other molecules contributing to
the function of target proteins, use is made of computer-
based modeling with the programs mentioned above.
Confirmation of essential contact sites in target proteins
is performed by mutational analysis and subsequent

25   identification of hydrophobic pockets located on the
protein surface in the vicinity of these contact sites.
        Computer modeling of "virtual" public compound
libraries for binding to these hydrophobic pockets and
testing of "best fit" compounds in in vitro (ELISA) and in

30   vivo (cell-based) assays for inhibition of target protein
function will allow determination of a structure-activity
relationship for compound classes.
        In addition, de novo compound libraries are

generated based on the information derived from the computer modeling described above using a combinatorial chemistry approach.

5   **Evaluation of candidate compounds**

Candidate compound are evaluated for cellular toxicity via commercially available service, such as MDS Pharma Services, USA. The evaluation of candidate compound efficacy in reducing polyp formation in APC^{min} mice

10 according to Su et al., (1992) Multiple intestinal neoplasia caused by a mutation in the murine homologue of the APC gene. Science 256: 668-670. The candidate compounds are also tested in other models predictive for colorectal cancer (e.g. Xenograft).

15

EXAMPLE 3

          <u>Quantitative PCR (Q-PCR)</u>

A panel of 20 different cancer cell lines derived

20 from lung (HOP-62, A549, EKVX), colon (DLD-1, Ls174T, HT29, HCT-116, SW480), breast (MDA-MB-435s, MCF7, T47D, MDA-MB-361, MDA-MB-231, MDA-MB-468) prostate (DU145, PC3), ovarian (IGROV-1, OVCAR-4) and melanoma (M14, SK-MEL-5) were lysed and total RNA was extracted using Trizol™ reagent

25 according to manufacturers' instructions. One microgram of RNA was reverse transcribed to generate the corresponding cDNA, which was used as the template for Q-PCR. The reverse transcription step was performed in 96-well plates using the TaqMan reverse transcription kit (Applied Biosystems)

30 according to the manufacturer's recommendations. The cDNA was quantified by the SyBR green method using a SyBR Green PCR master Mix kit (Applied Biosystems) according to the manufacturer's recommendations. For each reaction, 8 ng of

cDNA was used as a template and 300 nM of specific forward and reverse oligonucleotides added. Duplicate experiments were carried out using an Applied Biosytems 7000 SDS. Values were normalized according to the β-glucoronidase

5   (GUS) gene, which was measured as internal control. Oligonucleotides were designed using the Primer Express software (Applied Biosystems). These oligonucleotides were validated by Q-PCR experiments to obtain a quantitative measurement (quantification of serially diluted cDNA and

10   determination of PCR efficiency). The sequences of the oligonucleotides for internal control were as follows:

GUS: forward   5'-CCCGCGGTCGTGATGT-3'
           reverse   5'-TGAGCGATCACCATCTTCAAGT-3'

15

The sequences of the oligonucleotides used to probe the cDNA of the selected genes were:

EPHB2: forward 5'- TCTTCCTCATTGCTGTGGTTGT-3' (SEQ ID No. 13)

20   reverse 5'- TGTTGCAGCTTGTCCGTGTAC-3' (SEQ ID No. 14)
    GPX2:   forward 5'- CAGGGCCGTGCTGATTG-3' (SEQ ID No. 15)
           reverse 5'- CTCGTTGAGCTGGGTGAAGTC-3' (SEQ ID No. 16)
    RGMR: forward 5'- AGGAACGCTGGCACATTTTC-3' (SEQ ID No. 17)
           reverse 5'- TGAGTCCTAGACTGACAGACAAATCA-3' (SEQ ID No.

25 18)
    Tspan5:forward 5'- CTTCAATTGCACAGATTCCAATG-3' (SEQ ID No.
    19)
           reverse 5'- GGATCTTTAGTGCAGCAGGAGAA-3' (SEQ ID No.
    20)

30

                    Semi-Quantitative RT-PCR
              Total RNA was prepared from 18 different human
cancer cell lines of various origin using Fenezol™

according to the manufacturer's instructions (ActiveMotif, Belgium). First-strand cDNA was prepared from 1ug of total RNA using oligo dt (18) primers and MMLV RNA'se H minus point mutant reverse transcriptase (Promega). 2ul of the

5   25ul total first-strand cDNA was used as template in hot-start PCR's to amplify cDNA fragments corresponding to specific regions of the specified target gene products using primers spanning exon-intron boundaries (see below). GAPDH cDNA fragments were amplified from each cDNA sample

10   as an internal control for cDNA quality.

Specific primer combinations used to amplify the target gene cDNA fragments:

15   **Tspan5:**

Forward   5' – GCGAATTCGTGTCCGGGAAGCACTACAAG– 3' (SEQ
ID No. 21)
Reverse   5' – GCGAATTCGCCAGCTCGCCCTGACAGCTT –3' (SEQ
ID No. 22)

20

RGMR

Forward   5' – CACGCAGGTGACTGCCAACAG –3' (SEQ ID No.
23)
Reverse   5' – CGTCATCGATGCGTTCACTCA – 3' (SEQ ID No.
25   24)

EPHB3

Forward   5' – GGGTAACATCTGAGTTGGCGTGGA –3' (SEQ ID No.
25)
30   Reverse   5' – CATCGCCGTTGCAGTAGAGCTTG – 3' (SEQ ID No.
26)

GPR49

Forward   5' - CCTCAGTATGAACAACATCAGTCAG - 3' (SEQ ID

No. 27)

Reverse   5' - GCTGATGTGGTTAGCATCCAGAC - 3' (SEQ ID No.

5   28)

GPX2

Forward   5' - GCGAATTCGCTTTCATTGCCAAGTCCTTC - 3' (SEQ

ID No. 29)

10 Reverse   5' - GCCTCGAGCTATATGGCAACTTTAAGGAG - 3' (SEQ

ID No. 30)

    EXAMPLE 4

15

            CD44

            CD44 is a multistructural and multifunctional

cell surface receptor involved in cell-cell and cell-matrix

interactions, cell trafficking, presentation of chemokines

20 and growth factors to travelling cells and transmission of

growth signals. The extracellular matrix component (ECM)

hyaluronic acid (HA) is the principal ligand for CD44, but

other ligands include other ECM components and mucosal

addressin, serglycin, osteopontin and the class II

25 invariant chain Li. Multiple isoforms of CD44 exist (at

least 20) as a result of alternative splicing and

posttranslational modifications. The resulting spectrum of

products range in size from 85-23-Kda. The standard form of

CD44, which is the smallest isoform (CD44s/CD44H; 85Kda

30 following glycosylation) is the most abundant. It is widely

expressed, but is found at highest levels on haematopoietic

cells. Figure 10 shows a schematic representation of the

CD44 gene.

            Variant forms of CD44 (CD44v), generated by

alternative splicing around variable exons (V1-V10)
encoding extracellular domains are often more restricted in
their expression pattern. Expression of these variant forms
is often associated with activation of T-cells and some
5   studies indicate a correlation between CD44v expression and
cell proliferation. Many cancer cell types, as well as
their metastases, express high levels CD44 and in some
cancers there appears to be selection for high level
expression of particular variant forms of CD44. In colon
10 cancer, there is often selection for high level expression
of CD44v6 and CD44v9 (CD44v6/9 refers to all CD44 products
containing variant exon 6/9 and will include many CD44
variants composed of different combinations of constant and
variant exons). In these colon tumours this high level CD44
15 expression is driven by constitutive activation of the
TCF/β-catenin signalling pathway.

        The contribution of CD44 to carcinogenesis is
currently the focus of many studies. Over-expression of
CD44 in invasive colorectal cancers is associated with the
20 presence of metastases and with an unfavourable patient
prognosis. Interaction between CD44 and HA is proposed to
promote cell motility and sometimes tumour growth and
metastasis. Disruption of CD44 in metastatic mammary cancer
cells has also been shown to induce apoptosis, implying a
25 role for CD44 in counteracting programmed cell death. Other
studies indicate that inhibition of CD44 function in colon
cancers through administration of soluble CD44-Ig fusion
proteins or CD44 antisense RNA can retard tumour growth.
Targeting of CD44 variants differentially expressed on
30 colon cancer cells using specific antibodies may therefore
be a valid approach towards development of a safe and
effective therapy for colon cancer.

        TCF target gene status was confirmed by Northern

blot as shown in figure 2B. Figure 12 demonstrates that
CD44 expression requires TCF/β-catenin signalling. Thus,
inhibition of TCF/β-catenin signalling using dnTCF resulted
in a loss of CD44 expression on the cell-surface of colon

5 cancer cells. Furthermore it was demonstrated that CD44 was
lost in intestinal crypts following deletion of TCF-4 in
mice. Using pan-CD44 antibody for staining it was
demonstrated that CD44 is over-expressed in early
colorectal polyps compared to normal colon (Figure 13).

10   CD44v6 and CD44v9 were previously demonstrated to
be over-expressed in colon cancer (Wielenga et al.,Am. J.
Pathol., 1999, 154: 515-523). Other variant forms were also
more highly expressed in colon tumours than in normal
intestine.

15   Soluble CD44-Ig fusion proteins have been shown
to inhibit tumour formation by preventing binding of CD44
to its ligand(s). Accordingly, monoclonal antibodies are
generated. These antibodies are preferably prepared using
the full-length CD44 cDNA (containing variant exons V2-V10)

20 for the generations of immunogens, in order to maximize the
likelihood of obtaining monoclonal antibodies against
different CD44 variants. Purified His-tagged fusion
proteins generated from CHO cells and NS0 cell CD44
transfectants may be used to immunizetransgenic mice

25 expressing human immunoglobulins (such as Humab mice),
which will generate human antibodies against specified
antigens. Monoclonal antibodies specific for CD44 are
selected by screening NS0 transfectants using FACS
analysis. Functional assays are then performed to determine

30 the effects on apoptosis and proliferation/differentiation
of colon cancer cells in vitro and the efficacy of tumour
inhibition in mouse xenograft models.

        cDNA and protein sequences of CD44 and variants

which preferably are used according to the present
invention are given in Figure 17-19.

EXAMPLE 4

5

     <u>GPR49</u>

     Proteins in the large seven-transmembrane (TM), G
protein-coupled receptor (GPCR) superfamily are
functionally diverse and include receptors ranging from the

10 cAMP receptor in slime mold to mammalian neurotransmitters
and glycoprotein hormone receptors. GPR49 is most closely
related (35% homology at protein level) to a subgroup of
GPCRs that have a large N-terminal extracellular domain
containing leucine-rich repeats which are important for

15 interaction with large glycoprotein hormones, which leads
to cAMP production in target cells via activation of G-
proteins. GPR49 contains many more leucine-rich repeats
than other members of this family, indicating that it may
interact with larger (glycol) protein ligands. Specific

20 functions for GPR49 are currently unknown. Expression of
GPR49 is highest in muscle, placenta, spinal cord and
brain, but is present at lower levels in colon, small
intestine, bone marrow and adrenal gland (Hau et al. Mol.
Endocrinol. 1998, 12: 1830-1845).

25   It has been demonstrated that GPR49 is over-
expressed (3-fold) in almost half of hepatocellular
carcinomas compared to surrounding tissues (Yamamoto et al.
Hepatology, 2003, 37: 528-533). This high-level expression
correlated with the presence of β-catenin mutations in

30 hepatocellular cell-lines. Additionally, a strong GPR49
expression in colon cancer cells was demonstrated, which is
dependent upon TCF/β-catenin signalling.

     Moreover, the expression in colon carcinoma cells

was reduced 3-fold following inhibition of TCF/β-catenin
signalling (Table 1/Figure 2B).

        The GPR49 mRNA and protein sequences are given in
    Figure 20.

EXAMPLE 5

    EPHB4

        EPHB4 belongs to the EPH-related receptor
tyrosine kinase family, which has important roles in many
cellular processes, including neural development,
angiogenesis, vascular network assembly and proliferation.
Genetic studies using targeted mutagenesis in mice reveals
that EPHB4, together with the ligands ephrin-B1 and ephrin-
B2 is essential for the normal development of embryonic
vascular networks into arteries, veins and capillaries.
EPHB4 knockout mice die during embryogenesis, probably as a
result of failed cardiovascular development. EPHB2 and
EPHB3 are also expressed on veins and/or arteries and
combined loss of expression also leads to vascular defects,
although less pronounced than those seen for EPHB4. There
is evidence to suggest that EPH family members are
upregulated as blood vessels invade tumours, linking EPH
function with angiogenesis. EPHB4 is reported to be over-
expressed in ovarian cancer, endometrial tumors,
choriocarcinoma, teratocarcinoma and colon cancer.
Preliminary evidence indicates that this elevated
expression in colon cancer is a direct result of high level
TCF/β-catenin signalling, suggesting that EPHB4, as well as
EPHB2 and EPHB3 are TCF target genes. Given the likely
roles of EPHB4 (and other members) in tumour development,
it makes an attractive target for antibody-based therapies.

        The EPHB4 cDNA and protein sequence are given in

Figure 21.

EPHB4 is a receptor tyrosine kinase, class VIII
with a vestigial Ig-I domain, a single cysteine rich region
and two FNIII domains in the extracellular region. EPHB4 is
5   abundantly expressed in placenta and veins (but not
arteries) and in a wide range of primary tissues.

Northern analysis revealed EPHB4 to be a target
of TCF/β-catenin in colon carcinoma cells (Fig. 2B). In
accordance with this result, it was also shown that EPBH4
10 was highly expressed in intestinal crypts and over-
expressed in colorectal polyps (Figure 7).

EXAMPLE 6

15   GPX-2

GPX-2 is a member of the family of selenium-
dependent glutathione peroxidases, which are generally
thought to have an anti-oxidant function, protecting
tissues from reactive oxygen species. GPX-2 is the least
20 reliant of the family members on Selenium for expression,
due mainly to the stability of its mRNA during Selenium
depletion. It is highly expressed in the intestine and
liver and some reports indicate also a lower level
expression in the epithelium of the oesophagus. In the
25 mouse, GPX-2 maps close to a colon cancer susceptibility
locus and high expression levels correlate with resistance
to colon cancer. GPX-2 protein levels have also been shown
to increase in intestinal adenomas compared to adjacent
normal mucosa. This has led to speculation that cells may
30 upregulate GPX-2 (via transcriptional response to reactive
oxygen species) in the presence of reactive oxygen species
to protect against cancer by preventing the DNA damage that
would otherwise occur. According to the present invention

it was observed that GPX-2 expression is dependent upon
TCF/β-catenin signalling in colon cancer and recent studies
showing high-level expression in intestinal crypts and
early stage colon cancer indicate that it may in fact play
5   a role in regulating cell growth and differentiation, or
may serve to protect the developing cancer tissue from
oxidative stress.

The GPX2 mRNA sequence and protein sequence are
given in Figure 22.

10   GPX-2 is highly expressed in the intestine,
colon, stomach, liver and galbladder. Expression reduced 3-
fold in microarray experiments following inhibition of
TCF/β-catenin signalling (table 1). This was confirmed by
Northern blot analysis (Figure 2B).

15   Semi-quantitative and Q-PCR analysis of GPX-2
expression levels in a panel of human cancer cell-lines
demonstrated a significant expression in cancer cell-lines
of various origin, with particularly high levels evident in
colon carcinoma cell-lines LS174T, HT29, SW480, the lung

20   cancer cell-line A549 and the breast cancer cell-line
MDAMB351 (Figure 8 and Figure 9A).

EXAMPLE 7

25   RGMR

The Human Repulsive Guidance Molecule-Related
(RGMR) Gene (RGMR) is predicted to encode a 47KDa GPI-
anchored glycoprotein which is closely related to a family
of molecules designated repulsive guidance molecules (RGM)

30   present in humans, mice, chicken and Xenopus (Figure 29).
The RGM designation was assigned following the discovery of
the role chicken RGM has in axon guidance during nervous
system development. A murine orthologue of human RGMR

exists, indicating the existence of an RGM sub-family
(which was termed RGMR). No RGMR orthologue has been
identified in the chicken, suggesting that the RGMR genes
may have arisen via duplication of the RGM gene during
5   evolution.

Expression of human RGMR was found to be
dependent on active TCF/β-catenin signaling in colon cancer
cells (Table 1, Figure 2B).

It was previously shown that another family of
10 repulsive guidance proteins, the EPHB tyrosine kinase
receptors, are targets of TCF/β-catenin signaling and these
proteins are essential for the correct positioning of cell-
populations during development of the small intestine. RGMR
proteins may have similar roles to play in this
15 developmental process.

The hRGMR cDNA sequence and hRGMR protein
sequence are given in Figure 23. RGMR corresponds to image
sequence 376697 in Table 1.

This image sequence is 2kb downstream of the
20 actual predicted coding sequence of RGMR.

Semi-quantitative and Q-PCR analysis of RGMR
expression levels in a panel of human cancer cell-lines
demonstrated high-level expression in selected colon, lung
and prostate cell-lines, with highest levels evident in the
25 colon cancer cell-line LS174T (Figure 8 and 9C).

EXAMPLE 8

Tspan5
30 Tspan 5/NET4 is a member of the tetraspan
superfamily of proteins, which are characterized by four
transmembrane domains and two extracellular regions. Within
this superfamily is a sub-family of highly related proteins

referred to as NET proteins (for New Est Tetraspan), of
which there are currently more than 20 members. The
function of these NET proteins is largely unknown, although
it has been suggested that they may group together specific
5   cell-surface proteins including kinases to promote the
formation of stable functional signalling complexes.
General functions are considered to include regulation of
cell development, activation, growth and motility. The
limited data currently available suggests distinct roles
10 for individual tetraspan molecules, with some having
inhibitory roles in cancer and others apparently being
expressed at high levels in cancer cells. Tetraspan 5 is
one such family member that is expressed at high levels in
colon cancer, probably as a direct consequence of
15 constitutive TCF/β-catenin signalling. Expression levels
drop 3-fold following inhibition of TCF/β-catenin
signalling in colon cancer cell-lines (Table 1).

          The cDNA and protein sequence of Tspan5 are given
in Figure 24.
20   Figure 16 shows a schematic model of the domain
structure of Tspan5, demonstrating 4 transmembrane domains
and two large extracellular loops.

          The mouse homologue is mainly expressed in brain
tissue, but also in heart, kidney, testis and weakly in
25 liver. EST data on human Tspan5 indicate a similar
expression pattern.

          Semi-quantitative and Q-PCR analysis of Tspan5
expression levels in a panel of human cancer cell-lines
demonstrated high level expression in selected colon, lung
30 and prostate cancer cell-lines, with highest levels evident
in cancer cell-line HCT116 (Figure 8 and 9D).

EXAMPLE 9

Production of polyclonal antibodies

                    Polyclonal antibodies directed against selected

5   target surface-expressed target antigens (EPHB2, EPHB3,

Tspan5 and RGMR) were generated by immunization of rabbits

with gene-specific peptides predicted to be immunogenic and

adopt a conformation similar to that of the corresponding

region of the native protein (Software such as 'antigen

10   prediction' within the EMBOSS package of the UK HGMP

Resource Centre web-site was employed here). The presence

of antibodies directed against the target antigens was

confirmed by screening sera from immunized rabbits against

Cos7 cell transfectants expressing the appropriate full-

15   length target protein.

                    The following peptides were synthesized and

conjugated to BSA (bovine serum albumin) for rabbit

immunization:

20   EPHB2   Peptide 1:   H-YEKELSEYNATALKSPC-NH2

              Peptide 2:   H-PFSPQFASVNC-NH2

     EPHB3   Peptide 1:   H-PGSYKAKQGEGPC-NH2

              Peptide 2:   H-CQMNSVQLDGLPDARY-OH

25

     TSPAN5   Peptide 1:   H-CGYDARQKPEVDQQ-OH

              Peptide 2:   H-CKGVLSNISSITDLGGFD-OH

     RGMR   Peptide 1:   H-HSALEDVEALHPRKERC-NH2   .

30   Peptide 2:   H-CNYHSHAGAREHRRGD-OH

          The antibodies can be used as colon cancer

therapeutics via modulation of target protein function at

the cell-surface via inhibition of ligand binding or
inappropriate activation of downstream signaling cascades
in the cancer cells. Given the high-levels of expression of
these target proteins in colon cancer, these antibodies
5   will also be useful as diagnostic tools for colon cancer
incidence and progression.